ABSTRACT

A study is made of the function $\alpha_r(n)$ which denotes the digital sum of an integer n written in scale r. Explicit formulae for $\sum_{k<n} \alpha_r(k)$ are obtained and the results of previous papers concerning analytic approximations of this sum are given. Using the concept of normal numbers, some new results are obtained on asymptotic expressions for $\Sigma\alpha_r(k)$, k running through given normal sequences. The sum $\Sigma\alpha_r(n)$, r running over positive integers belonging to a given finite sequence, is considered and asymptotic estimates are obtained.

ON THE DISTRIBUTION OF SUMS OF DIGITS

by

John Reginald Trollope

Under the direction of

Dr. Leo Moser

Department of Mathematics

University of Alberta

A THESIS

SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

Edmonton, Alberta. September, 1955.

ACKNOWLEDGMENTS

I wish to extend my thanks to Dr. Leo Moser for his valuable assistance and suggestions. I would also like to express my appreciation to the National Research Council of Canada whose generous contributions made the preparation of this dissertation possible.

TABLE OF CONTENTS

INTRODUCTION

In this thesis we deal with problems that are related directly or indirectly with $\alpha_r(n)$, which denotes the digital sum of a positive integer n expressed in scale r. We present the relevant known results but some of the proofs given here are new. Some new problems are considered and the relevant results pertaining to these problems are presented.

In chapter I we consider a generalization of the conventional positional notation of integers. Theorems involving the digits in this representation are obtained and we show that in special cases the theorems reduce to known results.

In chapter II we return to the orthodox positional notation in scale r and consider the sum of $\alpha_r(n)$, n running over all positive integers less than some fixed number. The known results pertaining to this sum are given and some refinements are added.

Chapter III deals with sums of $\alpha_r(n)$, n running through normal sequences. We obtain asymptotic estimates for such sums.

In the concluding chapter we consider problems which are in some respects complementary to the problems dealt with in chapters II and III. Estimates are obtained for the sum of $\alpha_r(n)$, r running through a given sequence of positive integers.

CHAPTER I

ELEMENTARY FORMULAE RELATED TO DIGITAL SUMS

In the first part of this chapter we consider a generalization of the positional notation for the representation of integers, and derive explicit formulae for the sum of the digits. After considering some special cases of this generalization, we return to the conventional representation in scale r and prove some elementary theorems on divisibility.

It is well known that if r is any fixed integer greater than or equal to two, then, given any positive integer n, we can find integers a_k (k = o 1, 2, 3, ...) such that

$$n = \sum_{k=o}^{\infty} a_k r^k ,$$

where subject to the conditions

$$o \leq a_k < r \qquad (k=o,1, 2, 3, \ldots) ,$$

the integers a_k are uniquely determined, and in fact

$$a_k = \left[\frac{n}{r^k}\right] - r\left[\frac{n}{r^{k+1}}\right] .$$

I

(square brackets denoting integral parts according to the usual convention).

Also it may be shown that there exist unique integers a_k ($k=1, 2, 3, \ldots$) such that

$$n = \sum_{k=1}^{\infty} a_k k! ,$$

where the a_k are subject to the restrictions

$$0 \leq a_k < k + 1 , (k = 1, 2, 3, \ldots) .$$

This immediately suggests the following problem: what restrictions must be imposed on an arithmetic function $g(k)$ and on digits a_k in order that every integer n may be uniquely expressed in the form

$$n = \sum_{k=o}^{\infty} a_k g(k) .$$

We now obtain some sufficient conditions. Let $g(k)$ be an arithmetic function with the following properties:

(1.1) $$g(r+1) > g(r) , \quad (r=o,1, 2, 3, \ldots)$$

and

(1.2) $$g(0) = 1 ,$$

then we have,

THEOREM 1: Given any positive integer n there exist unique integers a_k,

$$0 \leq a_k < \frac{g(k+1)}{g(k)} , \quad (k=o, 1, 2, 3, \ldots),$$

such that

$$n = \sum_{k=o}^{\infty} a_k g(k) .$$

Proof: Since $g(r+1) > g(r)$, all $r \geq 0$, there exists an integer m such that

(1.3) $$g(m) \leq n < g(m+1).$$

By the division algorithm, there exist unique integers a_m, r_m such that

(1.4) $$n = a_m g(m) + r_m, \qquad 0 \leq r_m < g(m) ,$$

and clearly by (1.3)

(1.5) $$0 \leq a_m < \frac{g(m+1)}{g(m)}$$

Similarly there exist unique integers a_{m-1} and r_{m-1} such that

(1.6) $$r_m = a_{m-1}\, g(m-1) + r_{m-1}\ ,\quad 0 \le r_{m-1} < g(m-1)$$

where

(1.7) $$0 \le a_{m-1} < \frac{g(m)}{g(m-1)}\ .$$

Hence in general we have

(1.8) $$r_k = a_{k-1}\, g(k-1) + r_{k-1}\ ,\quad 0 \le r_{k-1} < g(k-1)$$

and

(1.9) $$0 \le a_{k-1} < \frac{g(k)}{g(k-1)}\ .$$

The r_k form a decreasing sequence of non-negative integers, therefore the process terminates, hence by iterating the expressions for the r_k we obtain

(1.10) $$n = \sum_{k=o}^{m} a_k g(k)\ ,$$

and by defining $a_k = 0$, $k > m$, we have the required representation.

If the $g(k)$ are restricted to functions satisfying

(1.1) $$g(k) \mid g(k+1)$$

then we have the following theorem:

THEOREM 2: If $n \geq 0$ is represented as in theorem 1 then,

$$a_k = \left[\frac{n}{g(k)}\right] - \frac{g(k+1)}{g(k)}\left[\frac{n}{g(k+1)}\right]$$

and

$$\sum_{k=o}^{\infty} a_k = n + \sum_{k=1}^{\infty} \left(1 - \frac{g(k)}{g(k-1)} \right) \left[\frac{n}{g(k)} \right] .$$

Proof: By theorem 1 we have

$$(1.12) \qquad \frac{n}{g(k)} = \frac{\sum_{i=k+1}^{\infty} a_i g(i)}{g(k)} + a_k + r \ , \ r < 1.$$

Therefore by (1.11)

$$\left[\frac{n}{g(k)} \right] = \frac{g(k+1)}{g(k)} \left[\frac{n}{g(k+1)} \right] + a_k \ ,$$

hence,

$$(1.13) \qquad a_k = \left[\frac{n}{g(k)} \right] - \frac{g(k+1)}{g(k)} \left[\frac{n}{g(k+1)} \right] .$$

Summing from k=0 to k = m, we obtain

$$(1.14) \quad \sum_{k=o}^{m} a_k = \sum_{k=1}^{m} \left(1 - \frac{g(k)}{g(k-1)} \right) \left[\frac{n}{g(k)} \right] - \frac{g(m+1)}{g(m)} \left[\frac{n}{g(m+1)} \right] + n \ ,$$

and theorem follows upon letting m increase without limit.

Corollary:

$$\sum_{k=o}^{\infty} (a_{2k+1} - a_{2k}) = \sum_{k=o}^{\infty} \left(1 + \frac{g(2k+1)}{g(2k)} \right) \left[\frac{n}{g(2k+1)} \right] -$$

$$- \sum_{k=1}^{\infty} \left(1 + \frac{g(2k)}{g(2k-1)}\right)\left[\frac{n}{g(2k)}\right] - n .$$

Proof: By theorem 2 we have,

$$(1.15) \qquad a_{2k+1} - a_{2k} = \left[\frac{n}{g(2k+1)}\right] - \frac{g(2k+2)}{g(2k+1)} \left[\frac{n}{g(2k+2)}\right]$$

$$- \left[\frac{n}{g(2k)}\right] + \frac{g(2k+1)}{g(2k)} \left[\frac{n}{g(2k+1)}\right] .$$

Hence for any $m > o$ we obtain

$$(1.16) \qquad \sum_{k=o}^{m} (a_{2k+1} - a_{2k}) = \sum_{k=o}^{m}\left(1 + \frac{g(2k+1)}{g(2k)}\right)\left[\frac{n}{g(2k+1)}\right] -$$

$$\sum_{k=1}^{m}\left(1 + \frac{g(2k)}{g(2k-1)}\right)\left[\frac{n}{g(2k)}\right] - n - \frac{g(2m+2)}{g(2m+1)} \left[\frac{n}{g(2m+2)}\right] .$$

Therefore by letting m approach infinity we obtain the required result.

We now consider some special cases of $g(k)$ satisfying (1.11). Let n be a positive integer represented in the scale of r, then if we denote its digital sum by $\alpha_r(n)$ we have:

THEOREM 3: (A. M. Legendre) [12]

$$\alpha_r(n) = (1 - r) \sum_{k=1}^{\infty} \left[\frac{n}{r^k}\right] + n .$$

Proof: In this case $g(k) = r^k$. The result then follows immediately from theorem 2.

THEOREM 4:

$$\sum_{k=o}^{\infty} (a_{2k+1} - a_{2k}) = (1 + r)\left(\sum_{k=o}^{\infty} \left[\frac{n}{r^{2k+1}}\right] - \cancel{\sum_{k=1}^{\infty} \frac{n}{r^{2k+1}}} - \sum_{k=1}^{\infty} \left[\frac{n}{r^{2k}}\right]\right) - n \; .$$

Proof: Replace $g(k)$ by r^k in the corollary to theorem 2.

THEOREM 5:

$$\sum_{k=o}^{\infty} a_{2k} = n - r \sum_{k=o}^{\infty} \left[\frac{n}{r^{2k+1}}\right] + \sum_{k=1}^{\infty} \left[\frac{n}{r^{2k}}\right]$$

and

$$\sum_{k=o}^{\infty} a_{2k+1} = \sum_{k=o}^{\infty} \left[\frac{n}{r^{2k+1}}\right] - r \sum_{k=1}^{\infty} \left[\frac{n}{r^{2k}}\right] \; .$$

Proof: From theorem 3 we have

$$(1.17) \quad \sum_{k=o}^{\infty} (a_{2k+1} + a_{2k}) = (1-r)\left(\sum_{k=o}^{\infty} \left[\frac{n}{r^{2k+1}}\right] + \sum_{k=1}^{\infty} \left[\frac{n}{r^{2k}}\right]\right) + n \; ,$$

hence the theorem is an immediate consequence of theorem 4.

We stated earlier in this chapter that every positive integer n may be expressed uniquely in the form

$$(1.18) \qquad n = \sum_{k=1}^{\infty} a_k \, k! \; .$$

Since by definition

$$(1.19) \qquad 0! = 1 = 1!$$

it follows that $g(k) = k!$ does not satisfy (1.1) for $k=1$, and it is evident that the representation of theorem 1 is not unique in digits a_0 and a_1.

However the difficulty can be surmounted as we need only to define a_0 as 0 and the validity of theorem 1 is restored.

THEOREM 6: If n is represented in the factorial scale, that is if,

$$n = \sum_{k=1}^{\infty} a_k k! \ , \quad 0 \le a_k < k+1 \ ,$$

then

$$\sum_{k=1}^{\infty} a_k = n - \sum_{k=2}^{\infty} (k-1) \left[\frac{n}{k!}\right]$$

Proof: Replace $g(k)$ by $k!$ in theorem 2.

We now consider a related problem. Let n be represented in scale r, that is $g(k) = r^k$, then we have by theorem 1 that

$$0 \le a_k < r \ ,$$

however it would be of interest to know how many digits a_k satisfy the inequality

(1.20) $$r - \mu \leq a_k < r \,, \quad 0 \leq \mu < r.$$

Denoting by ν the number of digits which satisfy (1.20) we have

THEOREM 7: (N. J. Fine) [9]

$$\nu = \sum_{k=1}^{\infty} \left[\frac{n + \mu r^{k-1}}{r^k} \right] - \frac{n - \alpha_r (n)}{r-1} \quad .$$

Proof: For a given fixed μ assume a_k satisfies (1.20), then

(1.21) $$n - r^{k+1} \left[\frac{n}{r^{k+1}} \right] \geq r^k (r - \mu).$$

Hence,

(1.22) $$\left[\frac{n}{r^{k+1}} + \frac{\mu}{r} \right] \geq 1 + \left[\frac{n}{r^{k+1}} \right] \quad .$$

However trivially by (1.20)

(1.23) $$\left[\frac{n}{r^{k+1}} + \frac{\mu}{r} \right] \leq 1 + \left[\frac{n}{r^{k+1}} \right] \quad ,$$

therefore we must have

(1.24) $$\left[\frac{n}{r^{k+1}} + \frac{\mu}{r} \right] = 1 + \left[\frac{n}{r^{k+1}} \right]$$

Using similar procedure it follows that $a_k < r - \mu$ implies

$$(1.25) \qquad \left[\frac{n}{r^k} + \frac{\mu}{r}\right] = \left[\frac{n}{r^{k+1}}\right]$$

Hence summing over all k we obtain

$$(1.26) \qquad \sum_{k=o}^{\infty} \left[\frac{n}{r^{k+1}} + \frac{\mu}{r}\right] = \nu + \sum_{k=o}^{\infty} \left[\frac{n}{r^{k+1}}\right] ,$$

and theorem follows immediately upon applying theorem 3 to the sum of the right hand side of (1.26). (For an alternate proof of above theorem see R. C. Buck [3]).

We now consider criteria for divisibility which are directly related to digital sums. Let n be a positive integer expressed in scale 10, then it can be shown that

$$n \equiv \alpha_{10}(n) \pmod 9$$

and that

$$n \equiv \alpha_{10}(n) \pmod 3$$

It is also true that

$$n \equiv \sum_{k=o}^{\infty} (a_{2k+1} - a_{2k}) \pmod{11} .$$

We now prove general results in scale r which reduce to the above if r = 10. Let n be a positive integer expressed in scale r, and let N be any positive integer such that

$$(1.27) \qquad (r-1) \equiv o \pmod{N}$$

Then we have

THEOREM 8:

$$n \equiv \alpha_r(n) \pmod{N}$$

PROOF: By theorem 3 we have

$$(1.29) \qquad \sum_{k=1}^{\infty} \left[\frac{n}{r^k}\right] = \frac{n - \alpha_r(n)}{r-1} \quad ,$$

hence, theorem follows since the left hand side of (1.29) is integral.

Alternatively let M be any positive integer such that

$$(1.30) \qquad (r + 1) \equiv o \pmod{M}$$

then we have the following theorem.

THEOREM 9:

$$n + \sum_{k=o}^{\infty} (a_{2k+1} - a_{2k}) \equiv o \pmod{M}$$

PROOF: By theorem 4 we have

$$(1.31) \sum_{k=o}^{\infty} \left[\frac{n}{r^{2k+1}}\right] - \sum_{k=1}^{\infty} \left[\frac{n}{r^{2k}}\right] = \frac{\sum_{m=o}^{\infty} (a_{2k+1} - a_{2k}) + n}{r + 1} \quad ,$$

and since the left hand side is integral the theorem follows.

In the remaining chapters we shall be concerned only with

the conventional positional notatio in scale $r \geq 2$.

Therefore for any fixed n, a_k will refer to the ceefficient of r^k in the representation of n in the scale of r.

CHAPTER II

EXPLICIT AND ASYMPTOTIC FORMULAS FOR THE AVERAGE SUM OF DIGITS

In the preceding chapter we obtained an explicit formula for the function

$$\alpha_r(n) = \sum_{k=o}^{\infty} a_k \quad ,$$

Namely

$$\alpha_r(n) = n - (r-1) \sum_{k=1}^{\infty} \left[\frac{n}{r^k} \right] .$$

However, a short examination of the eccentricities of this function is enough to dispel any hope of finding a simple analytical approximation to it, and, taking a second mean, we therefore transfer our attention to the sum

$$(2.1) \qquad A_r(n) = \sum_{m=0}^{n-1} \alpha_r(m) \quad ,$$

in which, we shall see, the more violent irregularities of the $\alpha_r(m)$ are smoothed out.

THEOREM 10: For $N = r^s$ $(s = 0, 1, 2 \ldots)$

$$A_r(N) = \frac{r - 1}{2 \log r} N \log N$$

This theorem can be proved by induction and by other elementary means. For the sake of variety we give here a

— proof based on a generating function.

PROOF: Consider the product

$$P_r(n) = \prod_{k=o}^{n} \left(1 + ax^{r^k} + a^2x^{2r^k} + \ldots + a^{r-1}x^{(r-1)r^k} \right).$$

Let m be any fixed integer $0 \leq m < r^{n+1}$, then in scale r we have,

$$m = a_p r^p + a_{p-1} r^{p-1} + \ldots + a_1 r + a_o$$

where $p \leq n$.

Since $x^{a_{p-i} r^{p-i}}$ $(i = 0, 1, 2 \ldots p)$ appears in one and only one factor of $P_r(n)$ it follows that x^m appears in the expansion of $P_r(n)$ and its coefficient is

$$a^{a_p + a_{p-1} + \ldots + a_o} = a^{\alpha_r(m)} .$$

Hence

$$(2.2) \qquad P_r(n) = \sum_{k=1}^{r^{n+1}-1} a^{\alpha_r(k)} x^k .$$

Differentiating $P_r(n)$ with respect to a we have,

$$(2.3) \quad \frac{d\, P_r(n)}{da} = \sum_{k=o}^{n} \frac{(x^{r^k} + 2ax^{2r^k} + \ldots + (r-1)a^{r-2}x^{(r-1)r^k})}{(1 + ax^{r^k} + \ldots + a^{r-1}x^{(r-1)r^k})} P_r(n)$$

From (2.2) it follows

$$\sum_{k=0}^{n} \frac{(x^{r^k} + 2ax^{2r^k} + \ldots + (r-1)a^{r-2}x^{(r-1)r^k})}{(1 + ax^{r^k} + \ldots + a^{r-1}x^{(r-1)r^k})} P_r(n) =$$

$$= \sum_{k=1}^{r^{n+1}-1} \alpha_r(k)\, a^{\alpha_r(k)-1} x^k$$

Letting $a = x = 1$ we have,

$$(2.4) \quad (n+1)\left(\frac{1+2+\ldots+(r-1)}{r}\right) \prod_{k=0}^{n} r = \sum_{k=1}^{r^{n+1}-1} \alpha_r(k)$$

$$(n+1)\frac{(r-1)}{2} r^{n+1} = \sum_{k=1}^{r^{n+1}-1} \alpha_r(k)$$

Hence if we denote r^{n+1} by N we have

$$(2.5) \quad A_r(N) = \frac{(r-1)}{2 \log r} N \log N \quad .$$

It is plausible from the above theorem that if an asymptotic approximation exists for $A^r(n)$ then

$$A_r(n) \sim \frac{(r-1)}{2 \log r} n \log n \quad .$$

The validity of this conjecture is shown in the following theorem.

THEOREM 11: (L. E. Bush) [4] .

$$A_r(n) \sim \frac{(r-1)}{2 \log r} n \log n \;. \; (n \to \infty).$$

PROOF: Consider all the integers from zero up to n-1 written in their natural order in scale r. The digits in the ith place from the right repeat themselves in periods of r^i numbers, each period consisting of r^{i-1} of each of the digits 0, 1, 2...(r-1). The last period will be complete if and only if r^i divides n. Let s_i be the sum of the digits in the ith place from the right in all the numbers from zero to n-1, then

$$(2.6) \qquad s_i \geq \left[\frac{n}{r^i}\right] \cdot \frac{r\,(r-1)}{2} \cdot r^{i-1}$$

$$> \frac{(r-1)}{2} \quad (n - r^i) \quad .$$

Also

$$(2.7) \qquad s_i < \left[\frac{n}{r^i} + 1\right] \cdot \frac{r\,(r-1)}{2} \cdot r^{i-1}$$

$$< \frac{(r-1)}{2} \quad (n + r^i)$$

But

$$\Lambda_r(n) = \sum_{i=1}^{k} s_i \qquad , r^{k-1} < n \leq r^k \quad .$$

Hence

$$(2.8) \qquad \Lambda_r(n) > \frac{1}{2}(r-1) \sum_{i=1}^{k} (n - r^i)$$

$$> \frac{1}{2}(r-1)kn - \frac{1}{2}(r-1)\, r^{k+1}$$

$$> \frac{1}{2}(r-1)\frac{n \log n}{\log r} - \frac{1}{2}(r-1)\, r^{k+1} .$$

Similarly

$$(2.9) \qquad A_r(n) < \frac{1}{2}(r-1) \sum_{i=1}^{k} (n + r^i)$$

$$< \frac{1}{2}(r-1)\, n \left(\frac{\log n}{\log r} + 1 \right) + \frac{1}{2}(r-1) r^{k+1} .$$

Therefore it follows from (2.8)

$$(2.10) \qquad \liminf_{n \to \infty} \frac{2\, A_r(n) \log r}{(r-1)\, n \log n} \geq 1 .$$

Also from (2.9) we have

$$(2.11) \qquad \limsup_{n \to \infty} \frac{2\, A_r(n) \log r}{(r-1) n \log n} \leq 1 .$$

Hence from (2.10) and (2.11) we have

$$A_r(n) \sim \frac{(r-1)}{2 \log r}\, n \log n .$$

The above theorem presents an asymptotic approximation for $A_r(n)$ but gives no estimate of the implied error. In the following theorems we present stronger results which give such estimates.

THEOREM 12 (R. Bellman and H. N. Shapiro) [1]

$$A_2(n) = \frac{n \log n}{2 \log 2} + O\,(n \log \log n) \quad .$$

The proof offered by the above authors is laborious and since stronger results for a general base are available, the proof of theorem 12 will be omitted.

THEOREM 13: (L. Mirsky) [13]

(i) $$A_r(n) = \frac{(r-1)}{2 \log r} \quad n \log n + O\,(n)$$

(ii) In the above $O(n)$ cannot be replaced by $o(n)$.

PROOF: We observe that if $k \geq o$ and $o \leq s \leq (r-1)$ then the representation of n in the scale of r contains the terms $s\ r^k$ if and only if n can be expressed in the form

(2.12) $$n = m\ r^{k+1} + \mu$$

where $m \geq o$, $s\ r^k \leq \mu < (s + 1)\ r^k$.

Hence if $f(n,k,\ s)$ is the number of positive integers not exceeding n whose representation in the scale of r contains the term $s\ r^k$ then

$$(2.13)\quad f(n,k,s) = r^{k}\left[\frac{n}{r^{k+1}}\right] + \sum_{\substack{\mu \le n - r^{k+1}\left[\frac{n}{r^{k+1}}\right] \\ s r^{k} \le \mu < (s+1)r^{k}}} 1$$

Clearly the second term on the left hand side is $O(r^{k})$, and removing square bracket we have

$$(2.14)\quad f(n,k,s) = \frac{n}{r} + O(r^{k}) \quad .$$

However

$$(2.15)\quad A_r(n) = \sum_{\substack{o \le k \le \log_r n \\ o \le s \le (r-1)}} s\, f(n,k,s) - \alpha_r(n)$$

$$= \sum_{\substack{o \le k \le \log_r n \\ o \le s \le r-1}} s\left(\frac{n}{r} + O(r^{k})\right) - \alpha_r(n)$$

Hence

$$(2.16)\quad A_r(n) = \frac{r(r-1)}{2}\left[\log_r n\right] \cdot \frac{n}{r} + O(r^{\log_r n}) - \alpha_r(n)$$

$$A_v(n) = \frac{(r-1)}{2 \log r} \quad n \log n + O(n) \quad ,$$

which complete the proof of (i)

To prove (ii) we consider the sequence

$$x_n = r^n (r+1)$$

By (2.13) we have

$$(2.17) \quad f(x_n, k, s) = r^k \left[\frac{x_n}{r^{k+1}} \right] + \sum_{\substack{\mu \le x_n - r^{k+1} \left[\frac{x_n}{r^{k+1}} \right] \\ sr^k \le \mu < (s+1) r^k}} 1$$

It follows immediately

$$(2.18) \quad f(x_n,k,s) = \begin{cases} \dfrac{x_n}{r} \quad , & k \le n-1 \\[2ex] r^n + \displaystyle\sum_{\substack{\mu \le r^n \\ sr^n \le \mu < (s+1) r^n}} 1 \quad , & k=n \\[4ex] 0 + \displaystyle\sum_{\substack{\mu \le x_n \\ sr^{n+1} \le \mu < (s+1) r^{n+1}}} 1 , & k=n+1 \end{cases}$$

Therefore from (2.15)

(2.19) $$A_r(x_n+1) = \sum_{0\le k\le n-1} \frac{sx_n}{r} + \left(\sum_{0\le s\le r-1} sr^n + 1\right) + (r^n + 1)$$

$0\le s\le r-1$

$$= \frac{(r-1)}{2} n\, x_n + \frac{r(r-1)}{2} r^n + r^n + 2$$

$$= \frac{(r-1)}{2} x_n \left[\log_r x_n - 1\right] + r^n \frac{r(r-1)}{2} + r^n + 2$$

$$= \frac{(r-1)}{2\log r} x_n \log x_n - C\, x_n + 2 \quad ,$$

where

(2.20) $$C = \frac{(r-1)}{2} \frac{\log(r+1)}{\log r} - \frac{r(r-1) + 2}{2r + 2} \neq 0$$

which completes the proof of (ii) .

In the above theorem no attempt was made to give bounds for the constant implied by the O. In the following theorems such bounds will be established.

Let $R_r(n)$ be defined by the equation

(2.21) $$A_r(n) = \frac{(r-1)}{2\log r} n \log n - R_r(n) \quad .$$

By theorem 10 it follows that if n is a power of r,

then,

$$(2.22) \qquad R_r(r^s) = o \qquad (s = 0, 1, 2 \ldots) \;.$$

We now consider bounds for $R_r(n)$ when n is arbitrary

THEOREM 14: (M. P. Drazin and J. S. Griffith) [8]

(i) For all relevant r, n

$$R_r(n) \geq 0$$

with equality if, and only if, n is a power of r.

(ii) For all relevant n

$$R_2(n) < \frac{n}{2} \;,$$

$$R_r(n) < \left(\frac{(r-1)\log(r-1)}{(r-2)\log r}\right) \frac{(r-1)n}{2} \;, \quad (r=3,4,5\ldots\ldots) \;.$$

The proof of theorem 14 will not be given here. However, using essentially the methods of Drazin and Griffith we shall prove the following refinement of theorem 14.

THEOREM 15: For all **positive** integers n and all integers $r \geq 2$

$$0 \leq R_r(n) \leq \frac{(r-1)}{2} \, n \;.$$

PROOF: Let $n = a\,r^s + b$ where $1 \leq a \leq r-1$, $s \geq o$ and $o \leq b < r^s$.

We see after a moments reflection, that

$$(2.23)\quad A_r(n) = a\,A_r(r^s) + r^s\,(1 + 2 + \ldots + (a-1)) + ba + A_r(b)$$

$$= a\,A_r(r^s) + \frac{a(a-1)}{2}\,r^s + ba + A_r(b)$$

By theorem 10.

$$(2.24)\quad A_r(n) = a\left(\frac{(r-1)}{2}\,r^s\,\log_r r^s\right) + \frac{a(a-1)}{2}\,r^s + ba + A_r(b)$$

But

$$\log(a\,r^s + b) = \log r^s + \log\,(a + b/r^s)\quad ,$$

and

$$ar^s = n - b\qquad ,$$

hence

$$(2.25)\quad A_r(n) = \frac{(r-1)}{2\log r}\,n\log n - \frac{(r-1)}{2\log r}\,\left(b\log r^s + n\log(a + b/r^s)\right)$$

$$+ \frac{a(a-1)r^s}{2} + ba + A_r(b)\quad .$$

Therefore by (2.21) we have,

$$(2.26)\quad R_r(n) = \frac{(r-1)}{2\log r}\,\left(n\log(a + b/r^s) + b\log r^s\right) - \frac{a(a-1)}{2}\,r^s$$

$$- b\,a - A_r(b)$$

$$R_r(n) = \frac{(r-1)}{2\log r}\left(n \log (a + b/r^s) + b \log r^s/b\right) - \frac{a(a-1)}{2} r^s$$

$$- b\,a + R_r(b) \quad .$$

To complete the proof we use induction on s, i.e., writing $n = ar^s + b$, we shall deduce the conclusion $0 \leq R_r(n) \leq \frac{(r-1)n}{2}$ from the hypotheses $o \leq R_r(b) \leq \frac{(r-1)}{2} b$.

We first prove that $o \leq R_r(n)$. Consider the case s = o (i.e. n < r); here we have

$$A_r(n) = (1+2+ \ldots + n-1) = \frac{n(n-1)}{2} \quad ,$$

and therefore

$$\text{(2.27)} \qquad R_r(n) = \frac{(r-1)}{2\log r}\, n \log n - \frac{n(n-1)}{2}$$

$$= \frac{n \log n}{2}\left(\frac{r-1}{\log r} - \frac{(n-1)}{\log n}\right) \geq o$$

Thus the result is proved for the case s = o. If we now suppose for a general n that $R_r(b) \geq o$ then by (2.26)

$$\text{(2.28)} \quad R_r(n) \geq \frac{(r-1)}{2\log r}\left(n \log (a+b/r^s) + b \log r^s/b\right) - \frac{a(a-1)}{2} r^s - ba.$$

Let $\theta = b/r^s$ and consider the function

$$(2.29)\quad f(a,\theta) = \frac{(r-1)r^s}{2\log r}\left((a+\theta)\log(a+\theta) - \theta\log\theta - \frac{2a\theta\log r}{r-1}\right)$$

Now if θ is taken as a continuous variable in the region $0 \le \theta \le 1$, we have

$$(2.30)\quad \frac{\partial f}{\partial\theta} = \frac{(r-1)r^s}{2\log r}\left(\log\frac{(a+\theta)}{\theta} - \frac{2a\log r}{r-1}\right) .$$

For small θ, $\frac{\partial f}{\partial\theta}$ is positive and can vanish at most once in the interval $0 \le \theta \le 1$. Hence the minimum value of $f(a,\theta)$ occurs either at $\theta = 0$ or $\theta = 1$, therefore

$$(2.31)\quad f(a,\theta) \ge \min\left(\frac{(r-1)r^s}{2\log r}\, a\log a\, ,\ \frac{(r-1)r^s}{2\log r}\,(a+1)\log(a+1) - \frac{2a\log r}{r-1}\right)$$

Hence

$$(2.32)\quad R_r(n) \ge \frac{ar^s(r-1)}{2}\left(\frac{\log a}{\log r} - \frac{(a-1)}{(r-1)}\right), \text{ or}$$

$$(2.33)\quad R_r(n) \ge \frac{(r-1)r^s}{2\log r}\left((a+1)\log(a+1) - \frac{2a\log r}{r-1} - \frac{a(a-1)}{r-1}\log r\right)$$

$$= \frac{(a+1)r^s(r-1)}{2}\left(\frac{\log(a+1)}{\log r} - \frac{a}{r-1}\right)$$

Since

$$g(a) = \left(\frac{\log a}{\log r} - \frac{a-1}{r-1}\right) = \frac{\log a}{(r-1)}\left(\frac{(r-1)}{\log r} - \frac{a-1}{\log a}\right)$$

is a *positive* function in the interval $1 < a < r$ and $g(1) = o$, the required result follows from (2.32) and (2.33).

To prove the second half of the inequality we again use induction on s. For $s = o$ we have by (2.27),

$$R_r(n) = \frac{n \log n}{2} \left(\frac{(r-1)}{\log r} - \frac{(n-1)}{\log n} \right)$$

$$= \frac{n(r-1)}{2} \left(\frac{\log n}{\log r} - \frac{(n-1)}{(r-1)} \right)$$

$$\leq \frac{n(r-1)}{2} \quad .$$

Hence proposition is true for $s = o$. Now for general n and assuming $R_r(b) \leq \frac{(r-1)}{2} b$ we have by 2.26,

$$(2.34) \quad R_r(n) \leq \frac{(r-1)r^s}{2 \log r} \left((a + \theta) \log (a + \theta) - \theta \log \theta \right)$$

$$- \frac{a(a-1)}{2} r^s - a r^s \theta + \frac{(r-1)b}{2} \quad .$$

From (2.30) it follows $R_r(n)$ has a maximum either at $\theta = 1$ or at $\theta = \xi < 1$, where ξ satisfies the equation

$$(2.35) \quad \log (a + \xi) - \log \xi - \frac{2\, a \log r}{r-1} = o \quad .$$

If the former case is true (i.e. maximum obtained when $\theta = 1$) we have from (2.34)

$$(2.36) \quad R_r(n) \le \frac{(r-1)r^s}{2 \log r} \left((a+1) \log (a+1) \right) - \frac{a(a+1)}{2} r^s + \frac{(r-1)b}{2}$$

$$= \frac{(r-1)}{2} \left(\frac{ar^s \log (a+1)}{\log r} + b \right) - \frac{r^s(r-1)}{2} \left(\frac{a(a+1)}{(r-1)} - \frac{\log(a+1)}{\log r} \right) .$$

We note that for $r \le 3$

$$\left(\frac{a(a+1)}{(r-1)} - \frac{\log(a+1)}{\log r} \right) \ge 0$$

Hence

$$R_r(n) \le \frac{(r-1)}{2} \left(\frac{ar^s \log (a+1)}{\log r} + b \right) \le \frac{n(r-1)}{2} .$$

To prove the inequality for arbitrary $r \ge 4$ we consider three cases, namely

(i) $a > \left[\sqrt{r} \right]$

(ii) $a = \left[\sqrt{r} \right]$

(iii) $a < \left[\sqrt{r} \right]$

For case (i) we have by (2.36)

$$(2.37) \quad R_r(n) \le \frac{(r-1)}{2} (ar^s + b) - \frac{r^s(r-1)}{2} \left(\frac{r}{r-1} - 1 \right)$$

$$\le \frac{(r-1)}{2} n$$

Similarly for case (ii) we have

$$(2.38)\quad R_r(n) \leq \frac{(r-1)}{2}\left(ar^s \log_r(\sqrt{r}+1) + b\right) \frac{-r^s(r-1)}{2}\left(\frac{a(a+1)}{(r-1)} - \log_r(\sqrt{r}+1)\right)$$

$$= \frac{(r-1)}{2}\left(\frac{ar^s}{2} + ar^s\log_r(1+1/\sqrt{r})+b\right) - \frac{r^s(r-1)}{2}\left(\frac{a(a+1)}{(r-1)}\right.$$

$$\left. - 1/2 - \log_r(1+1/\sqrt{r})\right)$$

$$= \frac{(r-1)}{2}\left(ar^s - \frac{ar^s}{2} + ar^s\log_r(1+1/\sqrt{r})+b\right) - \frac{r^s(r-1)}{2}\left(\frac{a(a+1)}{(r-1)}\right.$$

$$\left. - 1/2 - \log_r(1 + 1/\sqrt{r})\right)$$

$$= \frac{(r-1)n}{2} - \frac{r^s(r-1)}{2}\left(\frac{(a-1)}{2} + \frac{a(a+1)}{(r-1)} - (a+1)\log_r(1+1/\sqrt{r})\right).$$

Since $r \geq 4$ implies

$$(2.39)\qquad 1 + 1/\sqrt{r} < r^{1/3},$$ we have

$$(2.40)\qquad R_r(n) \leq \frac{(r-1)n}{2} - \frac{r^s(r-1)}{2}\left(\frac{(a-1)}{2} + \frac{a(a+1)}{(r-1)} - \frac{(a+1)}{3}\right)$$

$$\leq \frac{(r-1)n}{2} - \frac{r^s(r-1)}{2}\left(\frac{a}{6} + \frac{r-\sqrt{r}}{r-1} - \frac{5}{6}\right)$$

$$R_r(n) \leq \frac{(r-1)\,n}{2} - \frac{r^s(r-1)}{2}\left(\frac{a+1}{6} - \frac{(\sqrt{r}-1)}{r-1}\right)$$

$$\leq \frac{(r-1)\,n}{2} - \frac{r^s(r-1)}{2}\left(\frac{1}{2} - \frac{1}{3}\right)$$

$$\leq \frac{(r-1)\,n}{2} \quad .$$

Now we consider the final case. Since $a < \left[\sqrt{r}\right]$ implies $a + 1 \leq \left[\sqrt{r}\right]$, we have by (2.36)

$$(2.41) \quad R_r(n) \leq \frac{(r-1)}{2}\left(ar^s \log_r(\sqrt{r}) + b\right) - \frac{r^s(r-1)}{2}\left(\frac{2}{r-1} - \log_r(\sqrt{r})\right)$$

$$\leq \frac{(r-1)}{2}\,n - \frac{r^s(r-1)}{2}\left(\frac{a}{2} - \frac{1}{2} + \frac{2}{r-1}\right)$$

$$\leq \frac{(r-1)n}{2} \; .$$

We now have to consider the case when the maximum occurs at $\theta = \xi < 1$. By (2.34) we have

$$(2.42) \quad R_r(n) \leq \frac{(r-1)r^s}{2\log r}\left((a+\xi)\log(a+\xi) - \xi\log\xi\right)$$

$$- \frac{(a-1)r^s}{2} - ar^s\xi + \frac{(r-1)}{2}\,b$$

By (2.35) it follows.

$$(2.43)\quad R_r(n) \leq \frac{(r-1)r^s}{2\log r}\left((a+\xi)\log(a+\xi) - \xi\left(\log(a+\xi) - \frac{2a\log r}{r-1}\right)\right)$$

$$- \frac{a(a-1)r^s}{2} - a r^s \xi + \frac{(r-1)}{2} b$$

$$\leq \frac{(r-1)r^s}{2\log r} a\log(a+\xi) + ar^s\xi - \frac{a(a-1)r^s}{2} - ar^s\xi + \frac{(r-1)}{2} b$$

$$\doteq \frac{(r-1)}{2}\left(ar^s + b\right) - \frac{a(a-1)r^s}{2}$$

$$\leq \frac{(r-1)n}{2} .$$

This completes the proof.

The preceding results in this chapter have been devoted to obtaining analytical approximation for $R_r(n)$. We shall now obtain some explicit formulae for $R_r(n)$.

Let n be represented in scale 10 by

$$(2.44)\quad n = a_m 10^m + a_{m-1} 10^{m-1} + \dots + a_1 10 + a_o$$

Also define n_i , $i < m$ by

$$(2.45)\quad n_i = a_i 10^i + a_{i-1} 10^{i-1} + \dots + a_1 10 + a_o$$

Then we obtain,

THEOREM 16: (M. d'Ocagne) [14] For all positive integers n

$$A_{10}(n) = \sum_{i=1}^{m} a_i \left(10^{i-1} .5(a_i - 1 + 9i) + n_{i-1} \right) + \frac{a_o(a_o - 1)}{2} \quad .$$

The proof of theorem 12 will not be given however we shall prove a more general result from which the above theorem is obtained as a special case.

Let the representation of n in scale $r \geq 2$ be

$$(2.46) \qquad n = a_m r^m + a_{m-1} r^{m-1} + \ldots + a_1 r + a_o \quad ,$$

and define n_i , $i < m$ by

$$(2.47) \qquad n_i = a_i r^i + a_{i-1} r^{i-1} + \ldots + a_1 r + a_o \quad .$$

Then for all positive integers n we have

THEOREM 17:

$$A_r(n) = \sum_{i=1}^{m} a_i \left(\frac{r^i}{2} \left(i(r-1) + (a_i - 1) \right) + n_{i-1} \right) + a_o \frac{(a_o - 1)}{2}$$

PROOF: By (2.46), (2.47) and (2.23) we have

$$(2.48) \qquad A_r(n) = A_r(a_m r^m + n_{m-1})$$

$$= a_m A_r(r^m) + a_m \frac{(a_m - 1)}{2} r^m + a_m n_{m-1} + A_r(n_{m-1})$$

$$A_r(n) = a_m\left(\frac{r^m}{2}\quad m\,(r-1) + (a_m-1) + n_{m-1}\right) + A_r(n_{m-1})$$

Hence by iteration we obtain,

$$(2.49)\quad A_r(n) = \sum_{i=1}^{m} a_i\left(\frac{r^i}{2}\left(i(r-1) + (a_i-1)\right) + n_{i-1}\right) + A_r(n_o),$$

and since

$$A_r(n_o) = 1 + 2 + \ldots (a_o - 1) = a_o\frac{(a_o - 1)}{2}\quad ,$$

the proof is complete.

We now present an explicit formula for $A_r(n)$ which has the added advantage that it is readily converted to the asymptotic approximation of theorem 13.

THEOREM 18: If n and n_i are represented as in (2.46) and (2.47) respectively then,

$$A_r(n) = \frac{(r-1)}{2}\,mn - \frac{(r-1)}{2}\sum_{i=o}^{m-1} n_i + \sum_{i=o}^{m} a_i\left(\frac{(a_i-1)r^i}{2} + n_{i-1}\right).$$

PROOF: Consider the numbers from 1 to $n-1$ written in their natural order in scale r. Denote by S_{i-1} the sum of the digits in the i'th column from right. In the first column the digits repeat with period r, and the number of complete periods is $\left[\frac{n}{r}\right]$, hence

$$(2.50)\qquad S_o = \frac{r(r-1)}{2}\left[\frac{n}{r}\right] + a_o\,\frac{(a_o - 1)}{2}\quad .$$

In the second column the digits repeat in periods of length r^2, and sinse the sum of the digits in one period is $\frac{r^2(r-1)}{2}$, we have

$$(2.51)\qquad S_1 = \frac{r^2(r-1)}{2}\left[\frac{n}{r^2}\right] + r\left(1+2+,\ldots+(a_1-1)\right) + a_1 a_o$$

$$= \frac{r^2(r-1)}{2}\left[\frac{n}{r^2}\right] + r\,\frac{a_1(a_1-1)}{2} + a_1 n_o\quad .$$

In general it is easily seen tn t for $i \geq 1$

$$(2.52)\qquad S_i = \frac{r(r-1)}{2}\; r^i\left[\frac{n}{r^{i+1}}\right] + r^i\,\frac{a_i(a_i - 1)}{2} + a_i n_{i-1}$$

But by (2.46)

$$(2.53)\qquad A_r(n) = \sum_{i=o}^{m} S_i$$

hence

$$(2.54)\quad A_r(n) = \sum_{i=o}^{m}\left(\frac{r(r-1)}{2}\cdot r^i\cdot\left[\frac{n}{r^{i+1}}\right] + r^i\,\frac{a_i(a_i-1)}{2} + a_i n_{i-1}\right)$$

$$= \sum_{i=o}^{m-1}\frac{(r-1)}{2}\,r^{i+1}\left[\frac{n}{r^{i+1}}\right] + \sum_{i=o}^{m} a_i\left(\frac{(a_i-1)\,r^i}{2} + n_{i-1}\right).$$

However we note that

$$(2.55) \quad r^{i+1}\left[\frac{n}{r^{i+1}}\right] = n - n_i \quad ,$$

hence,

$$(2.56)\ A_r(n) = \frac{(r-1)}{2}\sum_{i=0}^{m-1}(n-n_i) + \sum_{i=0}^{m} a_i\left(\frac{(a_i-1)r^i}{2} + n_{i-1}\right)$$

$$= \frac{(r-1)mn}{2} - \frac{(r-1)}{2}\sum_{i=0}^{m-1} n_i + \sum_{i=0}^{m} a_i\left(\frac{(a_i-1)r^i}{2} + n_{i-1}\right).$$

which is the required expression.

We shall now show that theorem 18 may be derived directly from theorem 17. It might be noted however that the above proof is better in that it is more readily generalized.

By theorem 17 we have,

$$(2.57) \quad A_r(n) = \sum_{i=1}^{m} a_i\left(\frac{r^i}{2}\left(i(r-1) + (a_i-1)\right) + n_{i-1}\right) + \frac{a_o(a_o-1)}{2}$$

$$= \frac{(r-1)}{2}\sum_{i=1}^{m} i\, a_i r^i + \sum_{i=0}^{m} a_i\left(\frac{(a_i-1)r^i}{2} + n_{i-1}\right).$$

Considering the first term on the right hand side we note,

$$\sum_{i=1}^{m} i\, a_i\, r^i = m\left(a_m r^m + \dots + a_1 r + a_0\right) +$$

$$-\left(a_{m-1}\, r^m + \dots + a_0\right) +$$

$$\dots$$

$$-\left(a_1 r + a_0\right) +$$

$$-\ a_0 \quad .$$

Hence

$$(2.58) \qquad \sum_{i=1}^{m} i\, a_i r^i = m\, n - \sum_{i=0}^{m-1} n_i \quad ,$$

and result follows immediately from (2.57)

We shall now show that theorem 18 leads directly to the asymptotic formula of theorem 13. By theorem 18 we have

$$A_r(n) = \frac{(r-1)mn}{2} - \frac{(r-1)}{2} \sum_{i=0}^{m-1} n_i + \sum_{i=0}^{m} a_i \left(\frac{(a_i-1)r^i}{2} + n_{i-1} \right) .$$

Hence by (2.47)

$$(2.59) \quad A_r(n) = \frac{(r-1)mn}{2} + O\left(\frac{(r-1)}{2} \sum_{i=0}^{m-1} r^{i+1} + \sum_{i=0}^{m} \frac{(r-1)(r+1)r^i}{2} \right)$$

$$A_r(n) = \frac{(r-1)}{2} mn + O\left(r^{m+1} + (r+1)r^{m+1} \right)$$

$$= \frac{(r-1)}{2} \left[\log_r n \right] n + O(n)$$

$$= \frac{(r-1)}{2 \log r} n \log n + O(n) \quad .$$

We now consider a generalization of much of the above material. Let the representation of n in scale r be

$$n = a_m r^m + a_{m-1} r^{m-1} + \dots + a_1 r + a_o ,$$

and define:

$$\alpha_r^t (n) = \sum_{i=o}^{m} a_i^t ,$$

$$A_r^t (n) = \sum_{k<n} \alpha_r^t (k) , \text{ and}$$

$$\sigma_t (r) = \sum_{a=o}^{r-1} a^t , \qquad t \geq 0 \text{ , with } 0^0 \equiv 1 .$$

Then, we present the following results which reduce to previous theorems for t = 1.

THEOREM 19: If n and n_i are represented as in (2.46) and (2.47) respectively then,

$$A_r^t(n) = \frac{\sigma_t(r)}{r}\, mn - \frac{\sigma_t(r)}{r} \sum_{i=0}^{m-1} n_i + \sum_{i=0}^{m}\left(\sigma_t(a_i) r^i + a_i^t n_{i-1}\right) .$$

PROOF: Identical with that of theorem 18 with S_i replaced throughout by

$$(2.60) \quad S_i^t = \sigma_t(r)\, r^i \left[\frac{n}{r^{i+1}}\right] + r^i\, \sigma_t(a_i) + a_i^t n_{i-1} .$$

THEOREM 20:

$$A_r^t(n) = \sigma_t(r) \frac{n \log n}{r \log r} + O(n).$$

PROOF: For fixed t,

$$(2.61) \qquad \sigma_t(a_i) \leq \sigma_t(r) = O(1)$$

Hence

$$(2.62) \quad \frac{\sigma_t(r)}{r} \sum_{i=0}^{m} n_i - \sum_{i=0}^{m} \left(\sigma_t(a_i) r^i + a_i^t n_{i-1}\right) = O(r^{m+1}) = O(n).$$

Let $R_r^t(n)$ be defined by,

$$(2.63) \quad A_r^t(n) = \sigma_t(r) \frac{n \log n}{r \log r} - R_r^t(n).$$

Then, as a generalization of theorem 14 we have

THEOREM 21 (M. P. Drazin and J. S. Griffith) [8]

For all relevant r, t, n,

(i) $$R_r^t(n) \geq 0$$

with equality if and only if n is a power of r.

(ii) $$R_2^t(n) < \frac{\sigma_t(2)\, n}{2},$$

$$R_r^t(n) < \frac{\sigma_t(r)\, n}{r} \left(\frac{(r-1)}{(r-2)} \quad \frac{\log (r-1)}{\log r} \right)$$

PROOF: The procedure is similar to that used to prove theorem 15, therefore we shall omit the proof.

CHAPTER III

DIGITAL SUMS OF NORMAL SEQUENCES

The asymptotic values of the digital power sums $A_t^r(n)$ which we considered in the latter part of chapter II are in fact closely connected with the concept of the normality of sequences of digits as we shall now show.

Let x be a given real number in the interval $0 \leq x < 1$, and suppose it has the representation

$$(3.1) \qquad x = \sum_{i=1}^{\infty} \frac{a_i}{r^i} \qquad , r \geq 2$$

where

$$(3.2) \qquad 0 \leq a_i < r \qquad (i = 1,2,3 \ldots).$$

Given any digit $\underline{a}$ in this representation, let $N(r,x,m,a)$ denote the number of times that $\underline{a}$ occurs among the first m digits of this expansion of x. Clearly,

$$(3.3) \qquad \sum_{a=0}^{r-1} N(r,x,m,a) = m \quad .$$

The number x is said to be 'simply normal in the scale r', if

$$(3.4) \qquad \lim_{m \to \infty} \frac{N(r,x,m,a)}{m} = \frac{1}{r} \quad , (a = 1, 2, \ldots r-1).$$

Regarding r as fixed, let us express all the positive integers in the scale of r and write down their digital representations consecutively, with a 'decimal' point in

front, so as to define a real number x_r $(0 < x_r < 1)$.

Thus

$$x_{10} = .1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ \ldots \quad .$$

THEOREM 22: (Champernowne [5]) For each integer $r \geq 2$, x_r is simply normal in the scale r.

PROOF: The m'th digit in the expansion of x_r will have arisen as a digit in the expansion of some positive integer n, and clearly

$$(3.5) \qquad m = A_r^o(n) + O(\log_r n) \quad .$$

Hence by theorem 20

$$(3.6) \qquad m = \sigma_o(r) \frac{n \log n}{r \log r} + O(n) \quad ,$$

$$= \frac{n \log n}{\log r} + O(n) \quad .$$

Also

$$(3.7) \qquad N(r, x_r, m, a) = N(r, x_r, A_r^o(n), a) + O(\log n).$$

Now if the integers from 1 to n are written in a column, (as in the proof of theorem 18), it follows that if $F_i(n, a)$ is the number of times that a appears as a coefficient of r^i then

$$(3.8) \qquad f_i(a) = r^i \left[\frac{n}{r^{i+1}}\right] + O(r^i) \quad .$$

However we have identically

$$N(r, x_r, A_r^{\,o}(n), a) = \sum_{i=o}^{[\log_r n]} f_i(a) \quad ,$$

Hence

$$(3.9) \quad N(r, x_r, A_r^{\,o}(n), a) = \sum_{i=o}^{[\log_r n]} r^i \left[\frac{n}{r^{i+1}}\right] + O(r^i)$$

$$= \frac{n}{r} \left[\log_r n\right] + O(n)$$

$$= \frac{n \log n}{r \log r} + O(n) \quad .$$

Therefore by (3.6) (3.7) and (3.9) we have

$$(3.10) \quad \lim_{m \to \infty} \frac{N(r,x,m,a)}{m} = \lim_{n \to \infty} \frac{\left(\frac{n \log n}{r \log r}\right) + O(n)}{\left(\frac{n \log n}{\log r}\right) + O(n)}$$

$$= \frac{1}{r} \quad ,$$

and since a was an arbitrary digit the proof is complete.

In the above theorem we have used the properties of digital sums to prove normality. We shall now invert the procedure, that is, using the normality property of

sequences we shall obtain estimates for the sum of their digits.

Let $q_i (i = 1, 2, 3, \ldots)$ be a sequence of positive integers written in scale r, and let x be a real number (with $0 < x < 1$) defined by

$$(3.11) \qquad x = . q_1 \, q_2 \, q_3 \ldots ,$$

then we have the following lemma:

Lemma 1. If x is normal in scale r then,

$$\sum_{q \le q_k} \alpha_r(q) = \frac{(r-1)}{2} \sum_{q \le q_k} \log_r q + O(k) + o\Big(\sum_{q \le q_k} \log_r q \Big)$$

PROOF: Normality of x implies

$$(3.12) \qquad \lim_{m \to \infty} \frac{N(r,x,m,a)}{m} = \frac{1}{r} \quad , \ 0 \le a < r.$$

It is evident that (3.12) holds if m approaches infinity through any given subset of the positive integers. Hence let n_k denote the number of digits in the number

$$x_k = .q_1 \, q_2 \ldots q_k$$

then

$$(3.13) \qquad \lim_{m_k \to \infty} \frac{N(r,x,m_k, a)}{m_k} = \frac{1}{r} \quad .$$

However,

$$(3.14) \qquad m_k = \sum_{q \leq q_k} \left[\log_r q + 1 \right]$$

$$= \sum_{q \leq q_k} \log_r q + O(k) \quad .$$

Equation (3.13) can be written

$$(3.15) \quad N(r,x,m_k, a) = \frac{m_k}{r} + o(m_k)$$

Hence by (3.14) we have

$$(3.16) \qquad N(r,x,m_k,a) = \frac{1}{r} \sum_{q \leq q_k} \log_r q + O(k) + o(m_k)$$

It is evident that

$$(3.17) \qquad \sum_{q \leq q_k} \alpha_r(q) = \sum_{a < r} a N(r,x,m_k,a) \quad ,$$

therefore by (3.16), (3.14) and (3.17) we have,

$$(3.18) \quad \sum_{q \leq q_k} \alpha_r(q) = \frac{1}{r} \sum_{a < r} a \left(\sum_{q \leq q_k} \log_r q + O(k) + o(m_k) \right)$$

$$= \frac{(r-1)}{2} \sum_{q \leq q_k} \log_r q + O(k) + o\Big(\sum_{q \leq q_k} \log_r q \Big)$$

which completes the proof:

We now use lemma 1 to prove the following theorems.

THEOREM 23: For positive integral n and prime p

$$\sum_{p \leq n} \alpha_{10}(p) \sim \frac{9n}{2 \log 10}$$

PROOF: Copeland and Erdös [6] have shown that the decimal obtained from the sequence of primes, that is the number

(3.19) $\qquad x = .2\ 3\ 5\ 7\ 11 \ldots .$

is normal in scale 10, hence by lemma 1 we have

(3.20) $$\sum_{p \leq p_k} \alpha_{10}(p) = \frac{9}{2} \sum_{p \leq p_k} \log_{10} p + O(k) + o\Big(\sum_{p \leq p_k} \log_{10} p \Big) .$$

It is shown in 'Primzahlen' (Landau) [11] that

$$\sum_{p \leq m} \log p = m + O\Big(\frac{m}{\log^{q} m} \Big) \quad , \text{ for all } q.$$

Hence (3.20) becomes

(3.21) $$\sum_{p \leq p_k} \alpha_{10}(p) = \frac{9\, p_k}{2 \log 10} + O(k) + O\Big(\frac{p_k}{\log p_k} \Big) + o(p_k)$$

but by a weak form of the prime number theorem.

$$(3.22) \quad O(k) = O(\pi(p_k)) = o\left(\frac{p_k}{\log p_k}\right) = o(p_k)$$

Therefore

$$(3.23) \quad \sum_{p \le p_k} \alpha_{10}(p) = \frac{9p_k}{2 \log 10} + o(p_k) .$$

Let n be defined by

$$(3.24) \quad p_k \le n < p_{k+1}$$

Then,

$$(3.25) \quad \sum_{p \le n} \alpha_{10}(p) = \frac{9n}{2 \log 10} + O(n - p_k) + o(n) .$$

But by prime number theorem

$$(3.26) \quad \lim_{k \to \infty} \frac{p_{k+1}}{p_k} = 1 .$$

Hence,

$$(3.27) \quad O(n - p_k) = O(p_{k+1} - p_k) = o(p_k) \quad ,$$

$$= o(n) \quad ,$$

and the theorem follows from (3.25) .

THEOREM 24: For positive integral n and k

$$\sum_{k^2 \leq n} \alpha_{10}(k^2) \sim \frac{9}{2 \log 10} \sqrt{n} \log n \quad .$$

PROOF: It has been shown by Besicovitch [2] that the number obtained from the sequence of integral squares, that is,

$$x = .\ 1\ 4\ 9\ 16\ 25\ 36\ \ldots$$

is normal, hence by lemma 1 we have,

$$(3.28) \quad \sum_{k^2 \leq n^2} \alpha_{10}(k^2) = \frac{9}{2} \sum_{k^2 \leq n^2} \log_{10} k^2 + O(n) + o\left(\sum_{k^2 \leq n^2} \log_{10} k^2 \right)$$

$$= 9 \log_{10} n! + O(n) + o(\log n!) \quad .$$

However using a well known relationship, (see Hardy and Wright [10] , pp. 247 - 48) we have,

$$(3.29) \quad \log_{10} n! = \frac{n \log n}{\log 10} + O(n) \ .$$

Therefore by (3.28) and (3.29) it follows

$$(3.30) \quad \sum_{k^2 \leq n^2} \alpha_{10}(k^2) \sim \frac{9\ n \log n}{\log 10}$$

and on replacing n by $\sqrt{n}$ we obtain the theorem.

It has been shown by Davenport and Erdos [7], that if $f(n)$ is any polynomial in n that takes integral values when n is integral then the decimal x defined by

$$x = .\, f(1)\, f(2)\, f(3) \ldots \tag{3.31}$$

is normal, hence we have the following theorem,

THEOREM 25: If $f(n)$ is defined by

$$f(n) = a_o + a_1 n + \ldots + a_m n^m, \quad a_m > o,$$

then for any $N > 0$

$$\sum_{f(k)<N} \alpha_{10}(f(k)) \sim \frac{9}{2} \left(\frac{N}{a_m}\right)^{1/m} \log_{10} N \; .$$

PROOF: For any integer $N \geq a_o$ there exists an integer x such that

$$f(x) \leq N < f(x+1) \; . \tag{3.32}$$

Hence by lemma 1 we have,

$$\sum_{f(k)\leq N} \alpha_{10}(f(k)) = \sum_{k\leq x} \alpha_{10}(f(k))$$

$$\sim \frac{9}{2} \sum_{k \leq x} \log_{10} f(k) + O(x) \; . \tag{3.33}$$

However from the definition of $f(k)$ it follows that $\log_{10} f(x) \sim m \log_{10} x$. Hence.

$$(3.34) \qquad \sum_{f(k) \leq N} \alpha_{10}(f(k)) \sim \frac{9}{2} \sum_{f(k) \leq f(x)} m \log_{10} k \ ,$$

$$\sim \frac{9}{2}\, m\, x \log_{10} x \ .$$

From (3.32) it follows trivially that

$$(3.34) \qquad x = \left(\frac{N}{a_m}\right)^{1/m} + O(x^{\frac{m-1}{m}}) \ ,$$

hence we obtain the required result.

CHAPTER IV

SOME FURTHER RESULTS ON ANALYTIC APPROXIMATIONS OF SUMS OF DIGITAL SUMS.

In Chapters II and III we considered sums of $\alpha_r(k)$ in which r was fixed and k ran through various finite sequences of positive integers.

We now consider what is in a certain sense a complementary problem. Namely we investigate the behaviour of the sums of $\alpha_r(k)$ in which k is held fixed and r runs through some given sequence of positive integers.

Let q_i (i = 1, 2, 3,...) be an increasing sequence of positive integers, we define

(4.1) $$C(n) = \sum_{q \leq n} 1 ,$$

(4.2) $$\nu(n) = \sum_{q \leq n} (q-1) ,$$

and

(4.3) $$B_q(n) = \sum_{q \leq n} \alpha_q(n) .$$

We now prove the following lemma.

Lemma 2: For positive integral n

$$B_q(n) = n\,C(n) - \sum_{i \leq \sqrt{n}} \nu(n/i) + O\left(\sqrt{n}\,\nu(\sqrt{n}) + n\,C(\sqrt{n})\right) .$$

PROOF: Let $i \leq \sqrt{n} - 1$ be a positive integer, and if q is chosen such that

$$(4.4) \qquad \frac{n}{i+1} < q \leq \frac{n}{i}$$

then $q \geq \sqrt{n}$ and therefore by theorem 3 we have

$$(4.5) \qquad \frac{n - \alpha_q(n)}{q-1} = \left[\frac{n}{q}\right] .$$

But from (4.4)

$$\left[\frac{n}{q}\right] = i$$

Hence

$$(4.6) \qquad \alpha_q(n) = n - i(q-1) .$$

Summing over all the q that satisfy (4.4) we have by (4.3), (4.1) and (4.2),

$$(4.7) \qquad \sum_{q > n/i+1}^{q \leq n/i} \alpha_q(n) = \sum_{q > n/i+1}^{q \leq n/i} (n - i(q-1))$$

$$= n\Big(C(n/i) - C(n/i+1)\Big) - i\Big(\nu(n/i) - \nu(n/i+1)\Big)$$

However we have identically

$$B_q(n) - \sum_{q < \sqrt{n}} \alpha_q(n) = \sum_{q > \frac{n}{2}}^{q \leq n} \alpha_q(n) + \sum_{q > \frac{n}{3}}^{q \leq \frac{n}{2}} \alpha_q(n) + \cdots + \sum_{q < \sqrt{n}}^{q \geq \frac{n}{[\sqrt{n}-1]}} \alpha_q(n) .$$

Hence by (4.7),

$$(4.8) \qquad B_q(n) - \sum_{q < \sqrt{n}} \alpha_q(n) = n\Big\{\Big(C(n) - C(n/2)\Big) + \ldots + \Big(C(n/[\sqrt{n}-1]) - C(\sqrt{n})\Big)\Big\} -$$

$$- \Big\{\Big(\nu(n) - \nu(n/2)\Big) + \cdots + [\sqrt{n}-1]\Big(\nu\Big(\frac{n}{\sqrt{n}-1}\Big) - \nu(\sqrt{n})\Big)\Big\} .$$

$$= n\,C(n) - \sum_{i \leq \sqrt{n}} \nu\left(\frac{n}{i}\right) \quad + \quad O\left(n\,C(\sqrt{n}) + \sqrt{n}\,\nu(\sqrt{n})\right)$$

But, trivially

(4.19) $$\sum_{q < \sqrt{n}} \alpha_q(n) \leq n\,C(\sqrt{n}) \quad = O\left(n\,C(\sqrt{n})\right)$$

and lemma follows from (4.)

Lemma 1 is in itself too general to provide much information about $B_q(n)$, however it is useful in deriving asymptotic expressions for $B_q(n)$ once the sequence q_i is defined. By means of lemma 1 we now prove another lemma which leads directly to an asymptotic expression for $B_q(n)$ for some special sequences.

Let $R_1(n,r)$ and $R_2(n)$ be defined by the following equations

(4.1) $$C(n) = r\,C\left(\frac{n}{r}\right) + R_1(n,r) \quad ,$$

(4.11) $$\nu(n) = \frac{1}{2}\,n\,C(n) + R_2(n) \quad .$$

Then we have,

Lemma 3: For positive integral n and r

$$B_q(n) = \left(1 - \frac{\pi^2}{12}\right) n\,C(n) - \sum_{i \leq \sqrt{n}} \left(R_2\left(\frac{n}{i}\right) + \frac{n}{2}\frac{1}{i^2} R_1(n,i) \right)$$

$$+ \; O\left(\sqrt{n}\,\nu(\sqrt{n}) + n\,C(\sqrt{n})\right) \quad .$$

PROOF: By lemma 2, (4.11) and (4.1) ,

$$(4.12)\quad B_q(n) = n\,C(n) - \sum_{i \leq \sqrt{n}} \mathcal{V}\left(\tfrac{n}{i}\right) + O\left(\sqrt{n}\ \mathcal{V}(\sqrt{n}) + n\,C(\sqrt{n})\right)$$

$$= n\,C(n) - \sum_{i \leq \sqrt{n}} \left(\frac{1}{2}\,\frac{n}{i}\,C\left(\frac{n}{i}\right) + R_2\left(\frac{n}{i}\right)\right)$$

$$+ O\left(\sqrt{n}\ \mathcal{V}(\sqrt{n}) + n\,C(\sqrt{n})\right)$$

$$= nC(n)\left(1 - \sum_{i \leq \sqrt{n}} \frac{1}{i^2}\right) - \sum_{i \leq \sqrt{n}} R_2\left(\frac{n}{i}\right) - \frac{1}{2}\,n \sum_{i \leq \sqrt{n}} \frac{1}{i^2} R_1(n,i)$$

$$+ O\left(\sqrt{n}\ \mathcal{V}(n) + n\,C(\sqrt{n})\right) \quad .$$

However

$$(4.13)\quad \sum_{i \leq \sqrt{n}} \frac{1}{i^2} = \sum_{i=1}^{\infty} \frac{1}{i^2} - \sum_{i > \sqrt{n}}^{\infty} \frac{1}{i^2}$$

$$= \frac{\pi^2}{6} + O\left(\int_{\sqrt{n}}^{\infty} \frac{dx}{x^2}\right)$$

$$= \frac{\pi^2}{6} + O(1/\sqrt{n}) \quad ,$$

which completes the proof.

By means of lemma 3 we can easily prove the following two theorems.

THEOREM 26: For positive integral n and r.

$$\sum_{r \leq n} \alpha_r(n) = \left(1 - \frac{\pi^2}{12}\right) n^2 + O(n^{3/2})$$

PROOF: In this case we have

$$(4.14) \quad C(n) = \sum_{r \leq n} 1 = n$$

and

$$(4.15) \quad \mathcal{V}(n) = \sum_{r=1}^{n} (r-1) = \frac{n(n-1)}{2} .$$

Hence by (4.10) and (4.11) ,

$$(4.16) \quad R_1(n,r) = 0 ,$$

and

$$(4.17) \quad R_2(n) = -\frac{n}{2} .$$

Therefore

$$(4.18) \quad \sum_{i \leq \sqrt{n}} \left(R_2\left(\frac{n}{i}\right) + \frac{n}{2} \frac{1}{i^2} R_1(n,i)\right) + O\left(\sqrt{n}\, \mathcal{V}(\sqrt{n}) + nC(\sqrt{n})\right) =$$

$$= \frac{n}{2} \sum_{i \leq \sqrt{n}} \left(-\frac{1}{i}\right) + O(n^{3/2})$$

$$= O(n \log n + n^{3/2})$$

$$= O(n^{3/2}) \quad ,$$

and theorem follows immediately from lemma 3.

THEOREM 27: For positive integral n and prime p,

$$\sum_{p \leq n} \alpha_p(n) = \left(1 - \frac{\pi^2}{12}\right) \frac{n^2}{\log n} + O \frac{n^2}{\log^2 n} \quad .$$

PROOF: Here we have

$$(4.19) \qquad C(n) = \sum_{p \leq n} 1 = \pi(n)$$

and

$$(4.20) \qquad (n) = \sum_{p \leq n} (p-1) = \sum_{p \leq n} p - \pi(n).$$

It is shown in "Primzahlen" (E. Landau) [11] that

$$(4.21) \qquad \pi(n) = \frac{n}{\log n} + O\left(\frac{n}{\log^2 n}\right)$$

Also we require the following lemma:

Lemma 4

$$\sum_{p \leq n} p = \frac{n^2}{2 \log n} + O\left(\frac{n^2}{\log^2 n}\right)$$

PROOF: By partial summation,

$$(4.22) \qquad \sum_{p \leq n} p = \sum_{k=1}^{n} \Big(\pi(k) - \pi(k-1) \Big) \quad k$$

$$= \sum_{k=1}^{n-1} \pi(k) \Big(k - (k+1) \Big) + n\,\pi(n)$$

$$= (n+1)\,\pi(n) - \sum_{k=1}^{n} \pi(k)$$

However by (4.21)

$$(4.23) \qquad \sum_{k=1}^{n} \pi(k) = \sum_{k=2}^{n} \frac{k}{\log k} + O\left(\sum_{k=2}^{n} \frac{k}{\log^2 k} \right)$$

$$= \int_2^n \frac{u\,du}{\log u} + O\left(\frac{n}{\log n}\right) + O\left(\int_2^n \frac{u\,du}{\log^2 u} \right)$$

$$= \frac{n^2}{2 \log n} + O\left(\int_2^n \frac{u\,du}{\log^2 u} + \frac{n}{\log n} \right)$$

$$= \frac{n^2}{2 \log n} + O\left(\frac{n^2}{\log^2 n} \right) ,$$

Hence, the lemma follows immediately from (4.22) (4.23) and (4.21)

Now, it follows from (4.20), lemma 4 and (4.21) that,

$$(4.24) \qquad \nu(n) = \frac{1}{2}\, n \;\; \pi(n) + O\left(\frac{n^2}{\log^2 n} \right) ,$$

which by (4.11) implies

$$(4.25) \qquad R_2(n) = O\left(\frac{n^2}{\log^2 n}\right).$$

We must now obtain an estimate for $R_1(n, r)$. By (4.10) and (4.21) it is seen

$$(4.26) \qquad R_1(n,r) = \pi(n) - r\,\pi\left(\frac{n}{r}\right)$$

$$= \frac{n}{\log n} - r\left(\frac{n/r}{\log n/r}\right) + O\left(\frac{n}{\log^2 n/r}\right)$$

$$= n\left(\frac{\log n/r - \log n}{\log n \log n/r}\right) + O\left(\frac{n}{\log^2 n/r}\right)$$

$$= O\left(\frac{n \log r}{\log n \log n/r} + \frac{n}{\log^2 n/r}\right)$$

From lemma 3 we see we are only interested in $R(n, r)$ for $r \leq \sqrt{n}$, therefore

$$(4.27) \qquad R(n,r) = O\left(\frac{n \log r}{\log^2 n}\right), \quad r \leq \sqrt{n}.$$

By lemma 3 and equations (4.25), (4.27), (4.21) and (4.24) we have,

$$(4.29)\ \sum_{p\leq n} \alpha_p(n) - \left(1 - \frac{\pi^2}{12}\right) n\,\pi(n) = O\left(\sum_{i \leq \sqrt{n}} \frac{n^2}{\log^2 n}\left(\frac{1}{i^2} + \frac{\log i}{i^2}\right) + \frac{n^{3/2}}{\log n}\right)$$

$$= O\left(\frac{n^2}{\log^2 n} \sum_{i=0}^{\infty} \frac{(1+\log i)}{i^2}\right)$$

$$= O\left(\frac{n^2}{\log^2 n}\right),$$

which be (4.21) is equivilant to theorem 27.

It follows from lemma 3 that

$$(4.29)\qquad B_q(n) \sim \left(1 - \frac{\pi^2}{12}\right) n\,C(n)\ ,$$

for any sequence q_i for which

$$(4.30)\quad \sum_{i \leq \sqrt{n}} \left(R_2\left(\frac{n}{i}\right) + \frac{n}{2}\ \frac{R(n,i)}{i^2}\right) + \sqrt{n}\ \mathcal{V}(\sqrt{n}) + n\,C(\sqrt{n}) = o\left(n\,C(n)\right)$$

We shall now consider some sequences which satisfy this condition.

THEOREM 28: If q_i ($i = 1, 2, 3, \ldots$) is an arithmetic progression of positive integers having common difference d, then

$$B_q(n) \sim \left(1 - \frac{\pi^2}{12}\right) \cdot \frac{n^2}{d}$$

PROOF: Clearly,

$$C(n) = \left[\frac{n}{d}\right] \tag{4.31}$$

and

$$\mathcal{V}(n) = \left(\frac{2q_1 + d\left[\frac{n}{d}\right]}{2}\right)\left[\frac{n}{d}\right] + \left[\frac{n}{d}\right]\left(\frac{d}{2} + 1\right) \tag{4.32}$$

$$= \frac{n}{2}\left[\frac{n}{d}\right] + O(n)$$

Hence,

$$R_1(n,r) = O(r) \tag{4.33}$$

and

$$R_2(n) = O(n) \tag{4.34}$$

Denoting the left hand side of (4.30) by $F(n)$ we have by (4.33), (4.34), (4.32) and (4.31)

$$F(n) = O\left(n \log n\right) + O\left(n \log n\right) + O(n^{3/2}) + O(n^{3/2}) \tag{4.35}$$

$$= O(n^{3/2}) = o(n^{2})$$

and theorem follows from lemma 2.

BIBLIOGRAPHY

[1] Bellman, R. and Shapiro, H. N. On a problem in additive number theory. Ann. Math., Princeton (2), 49 (1948), 333 - 40.

[2] Besicovitch, A. S. The asymptotic distribution of the numerals in the decimal representation of the squares of natural numbers. Math. Zeit., 39 (1934), 146-56.

[3] Buck, R. C. Amer. Math. Monthly, 54 (1947), 487.

[4] Bush, L. E. An asymptotic formula for the average sum of digits of integers. Amer. Math. Monthly, 47 (1940), 154 - 6.

[5] Champernowne, D. G. The construction of decimals normal in the scale of ten. Jour. Lond. Math. Soc., 8 (1933), 254-60.

[6] Copeland, A. H. and Erdös, P. Note on normal numbers. Bull. Amer. Math. Soc., 52 (1946), 857 - 860.

[7] Davenport, H. and Erdös, P. Note on normal decimals. Can. Jour. Math. (1), 4 (1952), 58-63.

[8] Drazin, M. P. and Griffith, J. S. On the decimal representation of integers. Proc. Cam. Phil. Soc., (4), 48 (1952), 555 - 65.

[9] Fine, N. J. Amer. Math. Monthly, 53 (1946), 225.

[10] Hardy, G. H. and Wright, E. M. An introduction to the theory of numbers. (Oxford 1945).

[11] Landau, E. Handbuch der Lehre von der Verteilung der Primzahlen (Leipzig and Berlin 1908).

[12] Legendre, A. M. Theorie des nombres, ed 2. (1808).

[13] Mirsky, L. A theorem on representations of integers in scale r. Scripta Math., N. Y., 15 (1949) 11-12.

[14] d'Ocagne, M. Jorn. de sc. math. e ast. 7 (1886) 117-28.